

# Filtering Returns for Unspecified Biases in Priors when Testing Asset Pricing Theory

PETER BOSSAERTS

*California Institute of Technology and CEPR*

*First version received July 1999; final version accepted April 2002 (Eds.)*

Procedures are presented that allow the empiricist to estimate and test asset pricing models on limited-liability securities without the assumption that the historical payoff distribution provides a consistent estimate of the market's prior beliefs. The procedures effectively filter return data for unspecified historical biases in the market's priors. They do not involve explicit estimation of the market's priors, and hence, economize on parameters. The procedures derive from a new but simple property of Bayesian learning, namely: if the correct likelihood is used, the inverse posterior at the true parameter value forms a martingale process relative to the learner's information filtration augmented with the true parameter value. Application of this central result to tests of asset pricing models requires a deliberate selection bias. Hence, as a by-product, the article establishes that biased samples contain information with which to falsify an asset pricing model or estimate its parameters. These include samples subject to, *e.g.* survivorship bias or Peso problems.

## 1. INTRODUCTION

Standard tests of asset pricing models build on the premise that the empirical payoff distribution provides the right estimate of the market's historical priors. This is implicit in work on the efficient markets hypothesis (EMH; *e.g.* Fama (1970, 1991)), and explicit in the stationary, dynamic model of Lucas (1978), tested in Hansen and Singleton (1982), Mehra and Prescott (1985) and numerous other articles. The aim of this article is to derive procedures that allow one to test an asset pricing model without requiring that the historical *ex ante* beliefs of the market are directly related to the empirical payoff distribution. This is done for a class of securities to be referred to as limited-liability securities, distinguished because of a clear "default state", where the security pays nothing, and complementary states when the security pays a positive, possibly random amount. Examples include options (digitals, puts, calls) and equity. Because they can be viewed as the combination of a risk-free security and an option, corporate bonds and bank loans belong to this class as well.

The framework is very simple. Market prices are available for a security over several histories each comprised of a sequence of trading times  $t = 0, \dots, T$  and a terminal date  $T^* > T$ . The security's value at  $T^*$  is its liquidation value (if the security expires at  $T^*$ ) or its market price at  $T^*$  (if the security continues to live after  $T^*$ ). At any time prior to  $T^*$ , the deflated security price equals the conditional expectation of the deflated payout at  $T^*$ . The deflator reflects the market's discounting and attitudes towards risk. From the point of view of the empiricist, the deflator embodies the asset pricing model that (s)he is interested in testing and/or estimating its parameters. The conditional expectation is taken with respect to the markets' beliefs, which can be factored into a time-0 prior about the deflated payoff and a likelihood of all potential information flows given possible payoffs. It is standard to assume that the market's prior is correct, in the sense that it coincides with the distribution of payoffs over repetitions of histories.

It is also part of standard empirical analysis to assume that the likelihood is correct: the market knows the conditional frequency of all information patterns.

We relax these assumptions and require only that the likelihood function is correct. The market's prior beliefs of deflated time- $T^*$  payoffs may differ systematically from the recorded empirical frequency. We show how to manipulate price series such that the resulting return measures satisfy simple moment restrictions without having to estimate the market's prior at the beginning of each history, and indeed even without requiring that this prior remains the same across all histories. The manipulations involve the following: (i) if the price data originally come in the form of a long time series, rearranging them into a cross-section of shorter histories of length  $T^*$ ; (ii) selecting a subset of histories for analysis, picked on the basis of the price (payoff) level at the end (time  $T^*$ ), (iii) computing novel return measures. Effectively, the new return measures *filter* the data for any biases that may follow from systematic discrepancies between the market's prior and the recorded distribution of deflated payoffs across histories.

This includes situations where the final payoff is the same across all histories, yet the market thought that other outcomes were possible *ex ante*. Which may reflect mistaken beliefs on the part of the market, or it may simply be that the empiricist collected a biased sample. For instance, other outcomes were reasonably possible *ex ante*, but did not occur in the empiricist's sample. The latter has become known as the "Peso problem", referring to currency markets where devaluations could have been anticipated but did not happen until after the end of the time period under investigation. Another way in which the market's prior and the empirical distribution of payoffs may differ is through "survivorship bias": the sample only includes "winners" (limited-liability contracts with positive time- $T^*$  payout). This typically occurs in studies of equity markets, because stock that did not trade during at least a minimum time period is generally ignored (*e.g.* stock delisted for bankruptcy reasons).

In a recent paper, Lewellen and Shanken (2000) give examples of *ex post* return predictability when the market has to learn what the true empirical distribution of dividends was going to be, in which case there are obvious discrepancies between *ex ante* beliefs and *ex post* realizations. The present article develops statistics that can be used to filter return data for the biases induced by the learning in the Lewellen–Shanken paper. More structure on assets is imposed than in the Lewellen–Shanken paper (limited-liability securities); less structure on beliefs is required (neither the prior nor the likelihood function will have to be specified). There is an additional difference in the setup: in Lewellen and Shanken (2000), the market learns parameters of the distribution of dividends, the true values of which will never be revealed in finite time, whether to the market, or, more importantly, to the empiricist. In this article, the learning problem is formulated differently. The market learns about the value of the asset at some future date. Once this future date has arrived, the market observes the value of the asset unambiguously, as does the empiricist. If the asset has no expiration date, then this value is to be thought of as the market price (inclusive of any dividends). Such a formulation avoids estimation problems when the empiricist cannot observe what the market is learning about.

There are several motivations for attributing any systematic deviations between market beliefs and the joint empirical distribution of signals and payouts entirely to the market's prior at  $t = 0$ , while the likelihood function is assumed to be correct. First, standard mathematical analysis of Bayesian updating implicitly builds on this assumption. Second, knowing the likelihood function is tantamount to being able to read signals correctly. It is plausible that, while acquired in vastly different situations, this skill can be applied successfully in new situations, *i.e.* when pricing securities with a yet unknown payoff distribution.<sup>1</sup> Third, our approach leads

1. It may be useful to give a non-financial example to illustrate the meaning of the restrictions. Imagine an individual who is dropped one day at some random point on Earth where (s)he has never been before. (S)he is asked to

to elegant and easily verifiable restrictions on simple transformations (filters) of pricing data. Such results have yet to be obtained for the case where the market may be mistaken about the likelihood function. Fourth, beliefs have to be correct in some respect, because otherwise there is nothing that links prices to the empirical distribution of signals and payoffs. Without some restriction on beliefs, any price pattern that does not allow for arbitrage opportunities can be explained in terms of some set of beliefs.

The approach of the present paper should be reminiscent of some recent attempts to model portfolio decision making under model uncertainty, such as Cagetti, Hansen, Sargent and Williams (2000) and Maenhout (2001). In the latter papers, agents know the true model that generates future values, up to a parameter. The true model is referred to here as the likelihood function. Like in the present paper, agents are risk averse and deflate future payoffs using some known deflator. Cagetti *et al.* (2000) and Maenhout (2001) then derive optimal maximin portfolio decisions: for each portfolio choice, they determine expected future deflated payoffs by picking the parameter value that minimizes future deflated payoffs; subsequently, the portfolio decision is chosen that maximizes future expected deflated payoffs. Notice that agents do not formulate a prior over the unknown parameter and that there is no learning. This contrasts with the present paper, where agents are assumed to assign a prior to the unknown parameter, which is then updated using Bayes' law.

The filters all derive from a central property of Bayesian updating first proven here. It states that, when evaluated at the eventual outcome (actual parameter value that generated the series of signals at hand), the inverse of a Bayesian's posterior (one over her/his posterior) is a martingale. Because integrability problems would otherwise emerge, a technical assumption is imposed, referred to as the no early exclusion hypothesis (NEEH). Loosely stated, it means that no potential outcome is ever excluded prior to  $T^*$ . NEEH translates into a simple restriction on the information flow about limited-liability securities (the class of securities that we derive restrictions for), namely, that the default state never be revealed before  $T^*$ . Still, the procedures are not sensitive to small but non-negligible probabilities of early revelation of the default state, as would be the case with, *e.g.* equity contracts (where bankruptcy can generally be announced at any time).

The paper focuses on the econometric theory behind our estimation and testing procedures. A couple of empirical examples are included, for illustrative purposes only. Full-scale empirical implementations of the procedures can be found in Bondarenko and Bossaerts (2000), Bondarenko (2001), and Bossaerts and Hillion (2001). The remainder of the paper is therefore organized as follows. Section 2 introduces the mathematical framework and states the well-known equilibrium requirement that deflated prices should be conditional expectations of future deflated payoffs. Section 3 derives the main mathematical contribution, namely, that inverse posteriors form a martingale process (under NEEH). Section 4 applies this result to derive restrictions on the deflated prices of digital options and illustrates with data from the experimental IEM market. Section 5 studies general limited-liability securities. Daily, deflated CBOE S&P 500 index call options prices are used in an illustration. Section 6 elaborates on how to include preference parameters in the analysis. Section 7 discusses extensions to losers. Section 8 concludes.

predict the chance of rain the next day. Of course, because our individual visits the place for the first time, (s)he may not be well informed about the unconditional probability that it would rain the next day. Still, her/his experience with rain elsewhere should lead to a correct update of the prior. (S)he knows that rain is often preceded by a certain kind of clouds. These clouds are far less likely when the next day is rain-free. That is, (s)he knows the likelihood of types of news (type of clouds) given that it would rain the next day, or given that it would not rain the next day. With this information (signal and likelihood) (s)he updates and announces a prediction. Like our financial market, the individual holds a potentially biased prior, determined by her/his own particular experience elsewhere. Like our financial market, (s)he updates her prior using the correct likelihood function, because her/his experience elsewhere is fully relevant in the new situation.

## 2. PRICING ENVIRONMENT

We consider histories each of which consists of trading times  $t = 0, 1, \dots, T$  and a terminal date  $T^* > T$ , say  $T^* = T + 1$ . We are interested in studying the pricing of a security at all times prior to  $T^*$ , based on its payout  $\tilde{W}$  at  $T^*$ .  $\tilde{W}$  could be the security's liquidation value (if  $T^*$  is its expiration date) or the security's value at  $T^*$  (if it continues to live beyond  $T^*$ ). From the point of view of the market,  $\tilde{W}$  remains unknown prior to  $T^*$ . For simplicity, the security does not pay an intermediate dividend. Before each trading round, information is released about the likely value of  $\tilde{W}$ . The nature of this information can be left unspecified. To facilitate the analysis, however, it will be referred to as a "signal"  $S_t$ .

At  $t = 0$ , the market has some unspecified prior about  $\tilde{W}$ . As signals arrive at the marketplace, this prior is updated. The accumulated information from the signals prior to and including  $t$  will be denoted  $I_t$ . The collection  $\{I_t\}_{t=0}^{T^*}$ , therefore, has to be interpreted as the information filtration generated by the signals  $\{S_t\}_{t=0}^{T^*}$ . The time- $t$  market clearing price is assumed to be measurable with respect to  $I_t$ . Let  $\tilde{P}_t$  denote this price. Also, the final payoff,  $\tilde{W}$ , is measurable with respect to  $I_{T^*}$ .

As is standard in asset pricing theory, the market is assumed to set prices to satisfy a particular stochastic Euler equation. Let  $\tilde{R}_t$  denote the return on the asset:

$$\tilde{R}_t = \frac{\tilde{P}_t}{\tilde{P}_{t-1}}.$$

The return  $\tilde{R}_t$  solves the following stochastic Euler equation:

$$E^m[A_t \tilde{R}_t \mid I_{t-1}] = 1. \quad (1)$$

In this equation,  $A_t$  is a stochastic discount factor. Its form depends on the asset pricing model at hand. In Lucas' general equilibrium model (see Lucas, 1978),  $A_t$  equals the marginal rate of substitution of aggregate consumption over  $(t-1, t)$ :

$$A_t = \delta \left[ \frac{\partial u(c_{A,t})}{\partial c} \bigg/ \frac{\partial u(c_{A,t-1})}{\partial c} \right], \quad (2)$$

where  $c_{A,t}$  and  $c_{A,t-1}$  denote aggregate consumption at  $t$  and  $t-1$ , respectively, and  $u(\cdot)$  is the aggregate consumer's period utility function. Usually,  $A_t$  depends on some parameters that are unknown to the empiricist (in the above, these are the discount coefficient  $\delta$  and the parameters of the function  $u(\cdot)$ ), and hence, they will have to be estimated while testing the basic restriction in equation (1).<sup>2</sup> In Rubinstein's log-CAPM,  $A_t$  takes a particularly simple form

$$A_t = \frac{1}{\tilde{R}_t^M}, \quad (3)$$

2. As in the more recent literature on price behaviour under model uncertainty (Cagetti *et al.* (2000), Lewellen and Shanken (2000), Maenhout (2001)), it is being assumed throughout that the market knows the nature of the stochastic discount factor (*e.g.* in Lucas' model, it knows that the stochastic discount factor is its marginal rate of substitution of consumption over time, even if it may not know future consumption). Without this assumption, the market is schizophrenic, because it would ignore its own attitudes toward risk and how it plans to invest in the future, which is what the stochastic discount factor reflects. Note that the term "model uncertainty" in the literature does not refer to uncertainty about the asset pricing model (stochastic discount factor), but instead about the probabilistic model that generates future, exogenous payoffs (usually referred to as dividends), which the market is assumed to know only up to a parameter. Rather than working with a model of future payoffs given an unknown parameter, this paper will work with the likelihood of signals given the value at some future reference date. If the latter value is not an (exogenous) liquidating dividend, but instead the continuation price, it will depend not only on future exogenous random variables (such as dividends), but also on the market's own stochastic discount, which in turn may depend on other exogenous variables (*e.g.* total market dividend). This does not introduce any complication, because the market is assumed to know the nature of the stochastic discount factor. The approach of the present paper allows one to take the future payoff as the parameter that indexes model uncertainty, unlike in the aforementioned papers.

where  $\tilde{R}_t^M$  denotes the return on the market portfolio (supply of all assets in the economy). See Rubinstein (1976). As in the traditional, static capital asset pricing model (CAPM),  $A_t$  does not depend on any unknown parameters. The model is tested in, *e.g.* Bossaerts and Hillion (2001), one of the applications of the procedures presented in this article. It will also be used later in this paper to study the pricing of S&P 500 index options.

The expectation in the basic asset pricing restriction (1) has a superscript  $m$ . This is to make explicit that expectations are to be taken with respect to the market's subjective beliefs, which may differ from the objective measure that can be estimated consistently from the empirical distribution of realized  $A_t$ s and  $\tilde{R}_t$ s. In past econometric work, such discrepancies have not been allowed for. The advantage is that it makes verification of (1) simple when the world is stationary: one merely needs to test whether the sample version of the expectation on the L.H.S. of (1) is significantly different from 1. But the assumption may require implausibly accurate foresight. As a result, asset pricing models (choices of the stochastic discount factor  $A_t$ ) may be rejected even if they are otherwise perfectly valid. This paper is meant to explore to what extent discrepancies between market beliefs and the objective measure can easily be allowed for. The superscript  $m$  makes such discrepancies explicit.

Standard manipulations convert (1) into a form that is more useful for our purposes. Among other things, we will not insist that returns are stationary, and hence we won't need expressions that only involve returns. Also, we will deflate prices using a deflator constructed from the stochastic discount factor  $A_t$ . We will then work with deflated prices, making any reference to asset pricing models implicit. With deflated prices, the pricing restrictions will look *as if* the market is risk neutral and does not allow for time value of money; the market really may not be risk neutral and generally will discount for time value of money—all this is implicit in the deflating, however. The advantage is transparency. We can focus on market beliefs, the biases they generate with standard return measures, and construction of alternative return measures that can be restricted under the objective measure.

Write (1) in terms of prices:

$$\tilde{P}_{t-1} = E^m[A_t \tilde{P}_t \mid I_{t-1}]. \quad (4)$$

Apply this to the valuations at times  $T$  and  $T^*$ :

$$\tilde{P}_T = E^m[A_{T^*} \tilde{W} \mid I_T].$$

Since  $A_T, A_{T-1}, A_{T-2}, \dots, A_0$  are all in the market's information set at time  $T$ , we can also write

$$A_T A_{T-1} A_{T-2} \dots A_0 \tilde{P}_T = E^m[A_{T^*} A_T A_{T-1} A_{T-2} \dots A_0 \tilde{W} \mid I_T].$$

An analogous operation can be done at any prior point in time ( $t \leq T$ ):

$$A_{t-1} \dots A_0 \tilde{P}_{t-1} = E^m[A_t A_{t-1} \dots A_0 \tilde{P}_t \mid I_{t-1}].$$

Consequently, if we *scale* the price at  $t$  by the factor  $A_t A_{t-1} \dots A_0$ , and refer to the scaled price as  $P_t$ ,

$$P_t = A_t A_{t-1} \dots A_0 \tilde{P}_t, \quad (5)$$

then the fundamental asset pricing restriction in (4) becomes

$$P_{t-1} = E^m[P_t \mid I_{t-1}].$$

Iterating, one obtains

$$P_t = E^m[W \mid I_t], \quad (6)$$

with

$$W = A_T^* A_T A_{T-1} \dots A_0 \tilde{W}.$$

That is, the *scaled time- $t$  price* is the conditional expectation of the scaled payoff.

A comment about semantics. One generally refers to *deflated* prices and payoffs, instead of *scaled* prices and payoffs; the time- $t$  price deflator is  $1/[A_t A_{t-1} \dots A_0]$ ; the time- $T^*$  payoff deflator is  $1/[A_T^* A_T \dots A_0]$ . We will use the term deflation, since it has become standard in mathematical finance.

We will require that our description of beliefs and learning about the security's payoff applies to the deflated payoffs. That is, the market holds an arbitrary prior over the deflated payoff  $W$ , which factors into the security's own payoff  $\tilde{W}$  and the product of stochastic discount factors  $A_T^* A_T \dots A_0$ . Limited-liability securities pay zero in the default state, *i.e.*  $\tilde{W} = 0$  in the default state. But then  $W = 0$  as well. Also,  $A_t > 0$ , all  $t$ , so that  $W > 0$  when  $\tilde{W} > 0$ . Notice that limited-liability securities remain limited-liability securities even after deflating their payoff.

As mentioned before, it is standard in the empirical literature to assume that market beliefs about  $W$  coincide with the objective measure that the empiricist can estimate from realizations of  $W$ . We do not follow this tradition. Instead, we allow some arbitrariness in the market's time-0 prior belief about  $W$ . The nature of this arbitrariness is made more precise later on. Updating should take place according to the rules of conditional probability—Bayes' law. In fact, this is already implied by the conditional expectation in (6). In implementing Bayes' law, we require that the market uses the correct likelihood of the signals conditional on  $W$ . The meaning of "correct" should be clear: the likelihood function derives from the conditional probability with which signals have actually been generated over time. Hence, under the usual assumptions, an empiricist can estimate the likelihood consistently. The likelihood is implicit in historical sample moments that are computed conditional on the final payout  $W$ . Because the latter are easy to compute, we make them the focus of study. In other words, we attempt to derive restrictions on moments conditional on the final payout.

The restriction that the market uses the correct likelihood to update its beliefs reflects the idea that the market is able to correctly "read" the information it receives. Hence, we retain the core of what finance empiricists refer to as the EMH, namely, that "prices correctly reflect the available information" (see Fama, 1970, 1991). One could justify it by claiming that the market is able to rationally determine the link between eventual payouts and signals, even if this is not sufficient to determine the true marginal distribution of the final payouts.

Our study of the dynamics of securities prices in this setting will rely heavily on a simple property of Bayesian posteriors which seems to have gone unnoticed in the literature. Let us first turn to this property.

### 3. A RESTRICTION ON THE EVOLUTION OF BAYESIAN BELIEFS

Because the property we are about to introduce is of interest independent of the specific application in this article, it will be presented in as much generality as possible. Bossaerts (1999) illustrates that it can be used fruitfully to study the pricing of securities other than the limited-liability contracts that are the focus of this paper.

With one exception, we take the setup from the previous subsection. Time is denoted  $t = 0, 1, 2, \dots, T^*$ . At  $T^*$ , the value of a parameter  $V$  is announced. The Bayesian does not know the value of this parameter, but learns from a sequence of signals  $S_t$ ,  $t = 1, 2, \dots, T$  ( $T < T^*$ ), which form the information filtration  $\{I_t\}_{t=1}^T$ . Without loss of generality, take  $V$  to be continuous. Let  $\lambda_0(\cdot)$  denote the density corresponding to the time-0 prior about  $V$ . We allow the prior to be arbitrary.  $\lambda_0(\cdot)$  may even change across histories, which means that it may be different

for each sequence of signals drawn, as long as there is no dependence between the choice of  $\lambda_0(\cdot)$  and the drawing of signals from the true likelihood function. The notation does not make explicit the possible randomness of the prior, but is chosen to pursue simplicity. We do impose the usual restriction that the “true” probability that governs  $V$  is absolutely continuous with respect to the prior. This still allows the prior to put positive weight on outcomes that would never occur. The signal  $S_t$  may live in an abstract space, but, to simplify things, take it to be a continuous real random variable. We assume that  $V$  does not parametrize the range of  $S_t$ . The density of  $S_t$  given  $V$  (i.e. its likelihood), evaluated at  $s_t$  and  $v$ , is denoted<sup>3</sup>

$$l_t(s_t | v).$$

The subscript  $t$  allows the likelihood to change with the information prior to time  $t$ . In contrast with the prior, we assume that the likelihood is “correct”. The meaning of “correct” is: the likelihood function derives from the conditional probability with which signals are factually generated.

Let  $\lambda_t$  denote the time- $t$  posterior about  $V$ . It is based on the information in  $I_t$ .  $\lambda_t(v)$  can be obtained recursively using Bayes’ law:

$$\lambda_t(v) = \frac{l_t(s_t | v)\lambda_{t-1}(v)}{\int l_t(s_t | u)\lambda_{t-1}(u)du}. \quad (7)$$

The posterior is a function. Its value depends on the choice of its argument. For each history, choose the argument to be the realized value of  $V$ . This generates the random variable  $\lambda_t(V)$ . Because we assume that the prior is absolutely continuous with respect to the actual probability that governs  $V$ ,  $\lambda_0(V) > 0$ , and, hence,  $\lambda_t(V) > 0$ . Let us study changes in the posterior evaluated at  $V$ . In particular, what is the expected value of  $\lambda_t(V)$ , given  $I_{t-1}$ ? We insist on computing expectations under the “true” probability measure (the one that factually governs all the uncertainty). This way, we effectively study what an empiricist would observe in historical repetitions of the same environment. Let  $E[\cdot]$  denote the expectations operator corresponding to the “true” probability measure. As discussed in the previous section, we are looking for restrictions on moments conditional on the final outcome. The expectation conditional on the join of  $I_{t-1}$  and the information in  $V$  will be written as follows:

$$E[\cdot | I_{t-1}, V].$$

The following provides a first characterization of changes in Bayesian posteriors from the point of view of an outsider who knows the true parameter value.

**Lemma 1.**

$$E[\lambda_t(V) | I_{t-1}, V] \geq \lambda_{t-1}(V), \quad 0 < t \leq T.$$

The lemma simply states that the posterior evaluated at the eventual outcome is not expected to decrease. The significance of this lemma stems from the fact that it obtains independent of the market’s prior. For instance, the Bayesian may be excessively optimistic about certain values of  $V$ , which may lead one to expect decreases in the posterior in early periods (between  $t - 1$  and  $t$  when  $t \ll T^*$ ). Lemma 1 contradicts this.<sup>4</sup>

3. As is standard in probability theory, capitals are used to denote random variables (functions) and lowercase for draws.

4. Note that Lemma 1 does not imply that Bayesians eventually recover the true value of  $V$  or that Bayesian posteriors converge. This is consistent with known asymptotic results in the literature (see the counterexamples of non-convergence in Diaconis and Freedman (1986)).

Lemma 1 complements Doob's well-known result that Bayesian posteriors form a martingale process under the Bayesian's own beliefs (see Doob, 1948). Like Doob's result, the lemma characterizes changes of posteriors over any period in the learning process, no matter how small the previous signal history is. This contrasts with the asymptotic results about posteriors (such as asymptotic normality) which characterize the behaviour of posteriors only after large histories of signals. Unlike Doob's result, however, Lemma 1 characterizes changes in posteriors from the point of view of an outsider who knows the true parameter value, and not from the point of view of the Bayesian learner him/herself.

*Proof of Lemma 1.* From (7) (Bayes' law) and Jensen's inequality, it follows:

$$\begin{aligned}
 & E[\lambda_t(V) \mid I_{t-1}, V] \\
 &= E \left[ \frac{l_t(s_t \mid V)}{\int l_t(s_t \mid v) \lambda_{t-1}(v) dv} \mid I_{t-1}, V \right] \lambda_{t-1}(V) \\
 &\geq \frac{1}{E \left[ \frac{\int l_t(s_t \mid v) \lambda_{t-1}(v) dv}{l_t(s_t \mid V)} \mid I_{t-1}, V \right]} \lambda_{t-1}(V) \\
 &= \frac{1}{\int \frac{\int l_t(s_t \mid v) \lambda_{t-1}(v) dv}{l_t(s_t \mid V)} l_t(s_t \mid V) ds_t} \lambda_{t-1}(V) \\
 &= \lambda_{t-1}(V). \quad \parallel
 \end{aligned}$$

*Remark.* In this derivation,  $\lambda_{t-1}(V)$  is implicitly assumed to be in  $I_{t-1}$ . The generalization is straightforward: first condition on knowledge of the prior, then apply the law of iterated expectations. See Bossaerts (1996) for the general proof.

Consider now the *inverse* of the posterior,

$$\frac{1}{\lambda_t(V)}. \quad (8)$$

Is its behaviour restricted somehow? We are interested in conditional expectations of the inverse posterior. These would not be well defined if the inverse posterior were not integrable in the first place. That is likely to happen if  $\lambda_t(V)$  can become very small. The easiest way to avoid integrability problems is to assume that the time- $t$  posterior is bounded away from zero for  $v$  in a subset  $D$  of the range of  $V$ . We will then work with the truncated random variable

$$\frac{1}{\lambda_t(V)} 1_{\{V \in D\}},$$

where

$$1_{\{V \in D\}}$$

denotes the indicator function which takes the value of 1 if  $V \in D$  and 0 otherwise. With the convention that the product of zero and infinity equals zero ( $0\infty = 0$ ), the truncated random variable  $(1/\lambda_t(V))1_{\{V \in D\}}$  will always be integrable, and conditional expectations are well defined. Intuitively, one may take  $D$  to be the entire range of  $V$ , in which case 1 replaces  $1_{\{V \in D\}}$  in all what follows.

The minimum bound on  $\lambda_t(v)$  for  $v$  in  $D$  implies that the time- $t$  information never excludes the outcomes in  $D$ . Therefore, let us refer to the assumption as the "NEEH". Here is the formal statement.



**Assumption 1 (No Early Exclusion Hypothesis (NEEH)).** *Let  $D$  be a subset of the range of  $V$ . For all  $t$  in  $\{0, 1, \dots, T\}$ , there exists an  $\varepsilon > 0$  such that*

$$\lambda_t(v) \geq \varepsilon,$$

*for  $v \in D$ .*

Turning back to the question of restrictions on the dynamics of the inverse posterior, we have the following lemma, which will be the core mathematical result for the remainder of the paper.<sup>5</sup>

**Lemma 2.** *Let  $D$  be a subset of the range of  $V$  for which NEEH:*

$$E \left[ \frac{1}{\lambda_t(V)} 1_{\{V \in D\}} \mid I_{t-1}, V \right] = \frac{1}{\lambda_{t-1}(V)} 1_{\{V \in D\}},$$

*$t = 1, 2, \dots, T$ . Equivalently,<sup>6</sup> for  $v \in D$ :*

$$E \left[ \frac{1}{\lambda_t(v)} \mid I_{t-1}, V = v \right] = \frac{1}{\lambda_{t-1}(v)}.$$

Lemma 2 provides a fundamental property of Bayesian updating: the inverse of the posterior evaluated at the true parameter value forms a martingale process, provided (i) the Bayesian uses the correct likelihood function, (ii) the NEEH holds. The martingale result obtains with respect to the true probability measure, which complements an earlier result of Doob's, who showed that Bayesian posteriors form a martingale process relative to the beliefs expressed by the prior, as mentioned before.

Before we proceed to prove the lemma, it should be emphasized that the result is specified only for  $t = 1, 2, \dots, T$ . In particular,  $t < T^*$ . To make the lemma valid for  $t = T^*$  as well, we would have to extend the definition of NEEH to include  $T^*$ . But then the lemma becomes trivial. To see this, let us include  $T^*$  to derive the set  $D$  in NEEH.  $D$  is based on evaluation of the posterior at any time and across all possible histories. If we include  $T^*$ , then  $D$  would be the empty set. For when the outcome of  $V$  is announced at  $T^*$ ,  $\lambda_t(v) = 0$  for all  $v$  not equal to the announced outcome. This is true for any outcome of  $V$  that could possibly be announced, which implies that the minimal value of  $\lambda_t(v)$  is zero for *all*  $v$ . Hence,  $D$  is the empty set. With  $D$  equal to the empty set, the lemma is trivial ( $1_{\{V \in D\}} = 0$  always).

*Proof of Lemma 2.* From (7) (Bayes' law), it follows that

$$\begin{aligned} & E \left[ \frac{1}{\lambda_t(V)} 1_{\{V \in D\}} \mid I_{t-1}, V \right] \\ &= E \left[ \frac{\int l_t(s_t \mid v) \lambda_{t-1}(v) dv}{l_t(s_t \mid V)} \mid I_{t-1}, V \right] \frac{1}{\lambda_{t-1}(V)} 1_{\{V \in D\}} \\ &= \int \frac{\int l_t(s_t \mid v) \lambda_{t-1}(v) dv}{l_t(s_t \mid V)} l_t(s_t \mid V) ds_t \frac{1}{\lambda_{t-1}(V)} 1_{\{V \in D\}} \end{aligned}$$

5. Lones Smith independently proved a much weaker version of this result. He assumes, among other things, correct priors and independent signals. See Smith (1996).

6. By itself, the notation  $\lambda_t(v)$  does not make clear that we are only interested in the posterior evaluated at the *actual* outcome for  $V$  in each history.  $\lambda_t(v)$  would mislead one to think that we want the posterior evaluated at a  $v$  that is fixed across histories.

$$= \frac{1}{\lambda_{t-1}(V)} 1_{\{V \in D\}}. \quad \parallel$$

Let us now use this result in a study of the dynamics of the prices of limited-liability securities. First, consider the simplest ones, namely, digital options.

#### 4. PRICE DYNAMICS FOR DIGITAL OPTIONS

##### 4.1. Theory

With the aim of applying the results of the previous section, let  $V$  be binary:

$$V \in \{0, 1\}.$$

Use this to construct a simple security, with deflated payoff

$$W = V.$$

This is the payoff structure of a digital option. If  $W = 1$ , one states that the option expires “in the money”. If  $W = 0$ , the option matures “out of the money”. The latter is what we refer to as the default state.

We take the belief structure from the previous section. In particular, we assume that the market’s prior over  $V$  (hence,  $W$ ) is arbitrary, but that the market uses the correct likelihood function. The fundamental pricing equation (6) implies that the (deflated) price of the digital option equals the posterior evaluated at  $V = 1$ . With the notation of the previous section, this means

$$P_t = E^m[W \mid I_t] = E^m[V \mid I_t] = \lambda_t(1). \quad (9)$$

A simple application of Lemma 1 generates the following useful result.

**Theorem 1.** For  $t = 1, 2, \dots, T$ ,

$$E[P_t 1_{\{W=1\}} \mid I_{t-1}, W] \geq P_{t-1} 1_{\{W=1\}},$$

or equivalently,

$$E[P_t \mid I_{t-1}, W = 1] \geq P_{t-1}.$$

This theorem determines the nature of the selection bias that is introduced when conditioning on the “good” outcome ( $W = 1$ ). The impact is as expected: it causes an upward drift in prices. It should be emphasized that the theorem is *not* a simple consequence of the fact that the payoff equals 1 at  $T^*$ , and hence, that the price will ultimately have to increase to 1, in order to avoid arbitrage opportunities. Even if a trader knows that the payoff will be 1 at  $T^*$ , a priori it is not clear that (s)he will profit when forced to close out positions at any  $t$  before  $T^*$  ( $t \leq T < T^*$ ). Theorem 1 stipulates that (s)he may (it is not an arbitrage opportunity), and will on average.<sup>7</sup> As with Lemma 1, the significance of this theorem emerges when it is realized that it holds independent of the prior. Even if the market is excessively optimistic, there will be no temporary price decreases on average. The proof is an obvious application of Lemma 1.

The following result can be viewed as a prescription for correction of the selection bias in the previous theorem.

7. This puts into perspective recent attempts at modelling asset pricing anomalies by restricting the trading horizon of rational investors. See, e.g. De Long, Shleifer, Summers and Waldman (1990) and Liu and Longstaff (2000).

**Theorem 2.** *Provided the NEEH is satisfied for  $W = 1$ ,*

$$E \left[ \frac{1}{P_t} 1_{\{W=1\}} \mid I_{t-1}, W \right] = \frac{1}{P_{t-1}} 1_{\{W=1\}},$$

*$t = 1, 2, \dots, T$ . Analogously,*

$$E \left[ \frac{1}{P_t} \mid I_{t-1}, W = 1 \right] = \frac{1}{P_{t-1}}.$$

Theorem 2 states that inverse prices of winning digital options are a martingale under the true probability and relative to the market's information filtration, *when this filtration is augmented with the knowledge that  $W = 1$* . The proof is again a straightforward application of a previous lemma, Lemma 2.

Theorem 2 characterizes the dynamics of prices of digital options that expire in the money when the market uses the correct likelihood function in updating its beliefs, but may hold otherwise arbitrary priors. Hence, there is another way one can interpret Theorem 2. It proves that there is a hypothesis about market beliefs that can be falsified on samples that hold only “winners”. This hypothesis is weaker than the traditional EMH, because it requires that the market have correct beliefs only about the signals it receives conditional on the possible final outcomes, and not about the unconditional probability of the terminal payoff.

It deserves emphasis that Theorem 2 requires the NEEH *only* for the state  $W = 1$ . This means that the market must never be revealed at any point up to time  $T$  that the option is going to expire out of the money (in which case  $\lambda_t(1) = 0$ ). In contrast, it may be told early that the option will eventually expire in the money.

#### 4.2. Illustration

Theorem 1 implies that rates of returns of “winning” digital options will be non-negative on average. The rate of return is defined in the traditional way

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Theorem 1 states that

$$E[r_t \mid I_{t-1}, W = 1] \geq 0. \quad (10)$$

Theorem 2 then implies that a simple modification will make the resulting *modified* rates of return on winning digital options zero on average. The modification essentially filters the price series for the selection bias and any biases in the market's prior about the chances that  $W = 1$ . The modified rate of return, to be denoted  $x_t$ , is defined with the future price as basis:

$$x_t = \frac{P_t - P_{t-1}}{P_t}. \quad (11)$$

So, Theorem 2 predicts

$$E[x_t \mid I_{t-1}, W = 1] = 0. \quad (12)$$

Let us illustrate (10) and (12) on a small sample of historical digital option prices. While they are the building blocks of general equilibrium theory (where they are referred to as Arrow–Debreu securities), digital options are not widely traded. One exception is the Internet-based Iowa experimental market (IEM), organized by the University of Iowa. This is an electronic market in “winner-take-all” contracts (digital options). A few of those contracts derive their payoff from stock price changes of firms in the computer industry. Each month, a new set of

contracts is offered. Contract liquidation values are determined by changes in closing prices of the underlying stock measured from the third Friday of one month to the third Friday of the next month. Trade starts on Monday following the third Friday of the month. In our illustration, we will focus on the digital options that are written on common stock of Microsoft (MS). A comprehensive study of all IEM financial winner-take-all contracts can be found in Bondarenko and Bossaerts (2000).

Two MS digital options are traded. One, the “High” contract, pays a dollar when MS’s next month’s closing price is above a predetermined cut-off level. The other one, the “Low”, pays one dollar in the complementary state. The cut-off level is determined by the exercise price of the closest-at-the-money option written on MS and traded at the CBOE. We focus on daily closing prices, which are defined to be the last transaction price before midnight, or, if no transaction took place, the previous closing price.

Note that these securities are not digital options when one considers (i) real prices and payoffs (asset pricing theory applies only to real quantities), (ii) risk-adjusted (deflated) prices and payoffs. In real terms and after deflation, the securities either pay zero or some random, positive amount. To ensure that the MS winner-take-all contracts fit the assumptions of Section 4.1 so that they can be used as illustration, we therefore choose not to adjust for inflation and not to deflate prices (which means that we effectively assume risk neutrality and zero time value of money<sup>8</sup>). In the next section, the general case of limited-liability securities will be studied. These securities pay in real, deflated terms either zero or some positive, random amount. When inflation and adjustment for risk and time value of money is to be accounted for, only the general treatment covers MS winner-take-all contracts.

Table 1 displays cross-sectional averages of daily mean returns. That is, mean returns are computed for each contract (month) separately, and then averaged across contracts. For MS high, Table 1 shows that the average rate of return as usually computed (price change over lagged price) is 8.1%. With a  $t$ -statistic of 1.9, this is significant only at the 5% level.<sup>9</sup> This does provide some weak evidence against the traditional EMH, which would require the average rate of return to be equal to zero, assuming risk neutrality and zero time value of money. What is the evidence if we filter out biases in the market’s prior?

The price data are filtered for belief biases if we study *winning* MS digital options separately. Each month, one of the two digitals expires in the money, because they are complementary. So, we have 16 monthly series of daily rates of return on winning digitals. Across those 16 months, the average of the mean daily rate of return is 9.7%. With a  $t$ -statistic of 2.6, this is significant at the 1% level. Both the level and the significance level are higher than without the selection bias (we consider winning contracts only), which is to be expected from (10) (a restatement of Theorem 1).

Table 1 also displays the average *modified* rate of return. In each month, the MS digital option that expired in the money is retained, daily modified rates of return are computed from its price series, and their (time series) mean is computed. Subsequently, the mean modified rates of return across the 16 months are averaged, and the standard error is computed in the usual fashion. Table 1 documents that the average modified rate of return on winning MS digitals is –3.2%. With a  $t$ -statistic of –1.5, this number is insignificant. The insignificant –3.2% should

8. The zero time value of money may be justified by the fact that participants in IEM have to deposit \$500 to participate, and the deposit does not earn any interest, *whether one trades or not*. Hence, the opportunity cost to trading in IEM is leaving the \$500 on deposit, free of interest.

9. The significance levels in all tests of the paper are based on the  $t$ -distribution. There is no reason to suspect violations to the cross-sectional independence that would be required to validate the  $t$ -distribution, since cross-sections of non-overlapping time series are studied throughout.

TABLE 1  
*Securities prices in the Iowa experimental markets:  
 winner-take-all contracts written on MS common  
 stock*

	<i>N</i>	Average (%)
<i>All MS high contracts</i>		
Daily rate of return	16	8.1* (1.9)
<i>All MS winning contracts</i>		
Daily rate of return	16	9.7** (2.6)
Daily modified rate of return	16	−3.2 (−1.5)

*Remarks:* To compute the modified rate of return, the *end-of-period* price is used as basis; *N* is the number of time series (each is 1 month of end-of-day prices). The averages are computed as the cross-sectional average of the time series mean daily rate of return or modified rate of return. *t*-Statistics in parentheses. \*Significant at the 5% level, \*\*significant at the 1% level (both are one-sided significance levels).

be contrasted with the highly significant 9.7% average traditional rate of return on winning MS digitals. This confirms (12) (which is a restatement of Theorem 2).<sup>10</sup>

The example is included here only for illustrative purposes. A full analysis would require a more extensive data-set, inflation adjustment, as well as deflation to account for risk and time value of money. But that brings us outside the realm of digital options, and to the more general case, which we now consider.

## 5. PRICE DYNAMICS FOR GENERAL LIMITED-LIABILITY SECURITIES

### 5.1. Theory

Limited-liability securities have a peculiar payoff pattern. Whether deflated or not, they pay a positive amount except in a default state, where the security pays zero. So, letting  $W$  denote their deflated payoff, they have the following payoff structure:

$$W = \begin{cases} W^+ & (>0), \\ 0. \end{cases} \quad (13)$$

$W^+$  is a strictly positive random variable which determines the payoff if the security ends up outside the default state (in options jargon: the security expires “in the money”). Limited-liability securities can be considered “compound lotteries”: their payoff is determined by, first, a lottery  $V$  which determines whether the default state occurs ( $V$ ), and, second, a strictly positive lottery  $Y$  which determines the payoff if outside the default state. The payoff on a limited-liability security can therefore be expressed as

$$W = YV, \quad (14)$$

where  $V \in \{0, 1\}$  and  $Y > 0$ .

10. It would be interesting to develop statistical techniques to determine the significance level of the difference between the standard return from the entire data-set (8.1%), and the standard return (9.7%) as well as modified return (−3.2%) from the data-set of winners only. This is not straightforward, as it involves comparing non-nested hypotheses across overlapping but different data-sets. We leave such techniques for future research.

From our fundamental asset pricing formula (6), the pricing of limited-liability securities can readily be derived. For deflated prices and payoffs,

$$\begin{aligned} P &= E^m[W \mid I_t] \\ &= E^m[Y \mid I_t, V = 1]\lambda_t(1). \end{aligned} \quad (15)$$

As in the previous section, we will impose on  $V$  the belief structure that we introduced when deriving the fundamental property of Bayesian posteriors, Lemma 2: the market has an arbitrary prior over  $V$ , yet knows the likelihood of signals given  $V$ . As for the random variable  $Y$ , we will subject it to stricter assumptions.

**Assumption 2 (Correct Conditional Expectations (CCE)).** *The market's expectation about the payoff outside the default state is correct:*

$$E^m[Y \mid I_t, V = 1] = E[Y \mid I_t, V = 1].$$

Notice that only the conditional mean of the payoff is restricted. The market can have arbitrary priors in any other respect (e.g. conditional variance).

Because  $Y > 0$ ,  $E[Y \mid I_t, V] > 0$ . Just to avoid integrability problems when taking the inverse  $(1/E[Y \mid I_t, V])$ , assume that there exists  $\varepsilon > 0$  such that

$$E[Y \mid I_t, V] > \varepsilon.$$

We again study the effect of selecting “winners” only. That is, we condition on the knowledge that  $W > 0$ , or, equivalently,  $V = 1$ . We still need the NEEH for  $V = 1$ . The NEEH implies that the market never until  $T$  knows for sure that  $V = 0$ . Practically, this means that default must never be announced early. This no early exclusion of  $V = 1$  obviously limits the empirical scope of our results. For instance, in the case of equity (common stock), one can never exclude the possibility of early revelation of bankruptcy, the default state. Simulations, however, demonstrate that the results we have about the present are robust to the presence of a small probability of early revelation of the default state at any time up to  $T$ . See Bossaerts (1996).

The setup immediately leads to the following result.

**Theorem 3.** *Assume that the NEEH holds for  $V = 1$ . Assume CCE. For  $t = 1, 2, \dots, T$ ,*

$$E \left[ \frac{1}{P_t} 1_{\{W>0\}} \mid I_{t-1}, W \right] \geq \frac{1}{P_{t-1}} 1_{\{W>0\}}.$$

In the case of a digital option, we derived a martingale result when conditioning on the “good” outcome. For general limited-liability securities, we generate a submartingale result. Translated into returns (computed from deflated prices), this means that the *modified* rate of return on “winners” (limited-liability securities that mature “in the money”) is negative on average:

$$E[x_t 1_{\{W>0\}} \mid I_{t-1}, W] \leq 0.$$

The modified rate of return is defined as before:

$$x_t = \frac{P_t - P_{t-1}}{P_t}.$$

The above could be stated equivalently as follows:

$$E[x_t \mid I_{t-1}, W = w > 0] \leq 0. \quad (16)$$

*Proof of Theorem 3.* From (15), the assumption of CCE, Cauchy's inequality, Lemma 2 and the law of iterated expectations, it follows that

$$\begin{aligned}
& E \left[ \frac{1}{P_t} 1_{\{W>0\}} \middle| I_{t-1}, W \right] \\
&= E \left[ \frac{1}{E[Y \mid I_t, V=1] \lambda_t(1)} 1_{\{V=1\}} \middle| I_{t-1}, W \right] \\
&\geq E \left[ \frac{1}{E[Y \mid I_t, V=1]} \middle| I_{t-1}, W \right] E \left[ \frac{1}{\lambda_t(1)} 1_{\{V=1\}} \middle| I_{t-1}, W \right] \\
&\geq \frac{1}{E[Y \mid I_{t-1}, V=1]} \frac{1}{\lambda_{t-1}(1)} 1_{\{V=1\}} \\
&= \frac{1}{P_{t-1}} 1_{\{W>0\}}. \quad \parallel
\end{aligned}$$

Although empirically less attractive, we can restore the martingale restriction with an alternative assumption.

**Theorem 4.** Assume that the NEEH holds for  $V = 1$ . Assume CCE. If signals  $S_t$  ( $t = 1, 2, \dots, T$ ) and the outcome  $Y$  are conditionally independent, conditional on  $V = 1$ , then

$$E \left[ \frac{1}{P_t} 1_{\{W>0\}} \middle| I_{t-1}, W \right] = \frac{1}{P_{t-1}} 1_{\{W>0\}}.$$

Equivalently,

$$E \left[ \frac{1}{P_t} \middle| I_{t-1}, W = w > 0 \right] = \frac{1}{P_{t-1}}.$$

In terms of modified returns, this reads as follows:

$$E[x_t \mid I_{t-1}, W = w > 0] = 0.$$

The independence condition is objectionable. Take equity call options, for instance. Their payoff  $W$  derives from the price of the underlying equity. But the latter also signals the likelihood that the call option will expire in the money. Hence, signals and eventual payoffs are not independent. One case where the assumption of conditional independence does hold is the digital option. There,  $Y = 1$ , and, hence, trivially independent of any signal. Which effectively means that Theorem 2 can be considered to be a special case of Theorem 4.

*Proof of Theorem 4.* With the assumption of conditional independence of  $Y$  and the signals,

$$P_t = E[Y \mid V=1] \lambda_t(1).$$

Then, amend the proof of Theorem 3, as follows:

$$\begin{aligned}
& E \left[ \frac{1}{P_t} 1_{\{W>0\}} \middle| I_{t-1}, W \right] \\
&= E \left[ \frac{1}{E[Y \mid V=1] \lambda_t(1)} 1_{\{V=1\}} \middle| I_{t-1}, W \right] \\
&= E \left[ \frac{1}{E[Y \mid V=1]} \middle| I_{t-1}, W \right] E \left[ \frac{1}{\lambda_t(1)} 1_{\{V=1\}} \middle| I_{t-1}, W \right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{E[Y \mid V = 1]} \frac{1}{\lambda_{t-1}(1)} 1_{\{V=1\}} \\
&= \frac{1}{P_{t-1}} 1_{\{W>0\}}. \quad \parallel
\end{aligned}$$

There is a way to get a clean result without the objectionable independence assumption.

**Theorem 5.** Assume that the NEEH holds for  $W > 0$ . Assume CCE. For  $t = 1, 2, \dots, T$ ,

$$E \left[ \frac{W}{P_t} - \frac{W}{P_{t-1}} \mid I_{t-1} \right] = 0.$$

In words: *the change in ratio of the final payout over the price of a limited-liability security is zero on average.* Since  $W = 0$  when  $1_{\{W>0\}} = 0$ , the truncation with  $1_{\{W>0\}}$  is implicit. It was explicit in the previous results.

It is important to appreciate the difference in conditioning with the previous theorems. In Theorem 5, the *actual value of  $W$*  is *not* conditioned on. This has major empirical implications. Take Theorem 4, for instance. It predicts martingale behaviour for  $1/P_t$  when  $W > 0$ , and hence, imposes verifiable restrictions even on a single time series (provided  $W > 0$ ). In contrast, Theorem 5 is not a martingale result. Technically,  $W/P_{t-1}$  is not measurable in  $I_{t-1}$  ( $W$  doesn't become known until  $T^*$ , *i.e.* is measurable only in  $I_{T^*}$ ). Consequently, Theorem 5 restricts the price dynamics of only a *cross section* of winning limited-liability securities. It is not verifiable on a single price history.

Translated into returns computed from deflated prices, Theorem 5 implies that the *weighted modified rate of return* on limited-liability securities is zero on average. Mathematically,

$$E[x_t W \mid I_{t-1}] = 0. \quad (17)$$

Although superfluous because  $W = 0$  outside the event  $W > 0$ , one could have conditioned explicitly on  $W > 0$ :

$$E[x_t W \mid I_{t-1}, W > 0] = 0.$$

(This notation may erroneously give one the impression that the actual outcome of  $W$  is conditioned on, though!)

When contrasted with (16), the origin of (17) should be intuitive. In (16), the unweighted modified rate of return is non-positive on average (subject to some mild restrictions). The weighting in (17) restores the equality. Essentially, securities with a strongly negative modified return will also be those with a low  $W$ ; securities with a highly positive modified return will generally record a high  $W$ . The former are weighted less; the latter more. This turns the weak inequality into an equality. The fact that a simple weighting scheme corrects the inequality is a beautiful mathematical peculiarity.

*Proof of Theorem 5.* The main ingredients to this proof are (i) the law of iterated expectations, (ii) Lemma 2:

$$\begin{aligned}
&E \left[ \frac{W}{P_t} - \frac{W}{P_{t-1}} \mid I_{t-1} \right] \\
&= E \left[ \frac{YV}{E[Y \mid I_t, V = 1]\lambda_t(1)} - \frac{YV}{E[Y \mid I_{t-1}, V = 1]\lambda_{t-1}(1)} \mid I_{t-1} \right] \\
&= E \left[ E \left[ \frac{YV}{E[Y \mid I_t, V = 1]\lambda_t(1)} - \frac{YV}{E[Y \mid I_{t-1}, V = 1]\lambda_{t-1}(1)} \mid I_t, V \right] \mid I_{t-1} \right]
\end{aligned}$$



$$\begin{aligned}
&= E \left[ \frac{E[Y | I_t, V]V}{E[Y | I_t, V = 1]\lambda_t(1)} - \frac{E[Y | I_t, V]V}{E[Y | I_{t-1}, V = 1]\lambda_{t-1}(1)} \middle| I_{t-1} \right] \\
&= E \left[ \frac{1}{\lambda_t(1)} 1_{\{V=1\}} - \frac{1}{\lambda_{t-1}(1)} 1_{\{V=1\}} \middle| I_{t-1} \right] \\
&= 0. \quad \parallel
\end{aligned}$$

## 5.2. Illustration

To illustrate the theoretical developments, let us look at a sample of price histories of the S&P 500 index option contract (coded SPX) traded at the CBOE. Four-week series of daily prices of call options are investigated, covering the years 1991–1995. At the start of each series, we picked the option that was closest to the money and 5 weeks from expiration. The prices were estimated from closing prices of actually traded put and call options, using the smoothing technique developed in Bondarenko (1998) (constrained convex least squares regression). The technique filters the data for bid-ask bounce and violations of simple arbitrage bounds. Prices and payoffs are converted from nominal to real quantities on the basis of the U.S. consumer price index (CPI) numbers provided by the Federal Reserve Bank of Saint Louis. CPI numbers are monthly; linear interpolation is used to convert the monthly frequency to daily frequency.

Let us consider a simple but fully dynamic asset pricing model, namely, Rubinstein's log-CAPM. See Rubinstein (1976). The model generates the Black–Scholes call option pricing model as a special case. We incorporate Rubinstein's log-CAPM through deflation of (real) prices and payoffs before computing return measures, using the formulation in (5). The stochastic discount factor equals the inverse (real) return on the market portfolio. Hence, the scaling factor used to compute deflated prices and payoffs is the inverse of the cumulative (real) return on the market portfolio. See equation (3). We take the S&P 500 index as a proxy of the market portfolio. As a model for risk adjustment, Rubinstein's model comes in handy, because it generates a deflator that is parameter-free. That is, we can test the asset pricing model and our belief model without having to estimate any parameters.

Before we look at the results, we should remind ourselves of a standard pricing anomaly in options markets, namely, the smile effect, and discuss its relationship with potential results from the exercises we are about to perform. The smile effect concerns the pricing of out-of-the-money options relative to in-the-money options. Evidently, out-of-the-money options are far more expensive compared with at-the-money options when evaluated against a standard option pricing model, say, Black–Scholes. The mark-up increases with the extent to which the option is out of the money.

In the statistical exercises we are about to perform, we ask a related but different question: is there a smile effect even in realized, risk-adjusted real option returns? That is, does the risk-adjusted performance of options decrease as the strike price increases, effectively revealing that the more an option is out of the money, the more the option is overpriced?

If we find such a smile effect in the realized risk-adjusted returns, there could be two explanations. First, our model of risk adjustment may differ from the one used by the market. We posit that the market adjusts for risk as in Rubinstein's log-CAPM, which may be wrong. The second explanation is very different. It could be that the market does price risk in accordance with Rubinstein's log-CAPM, but its *ex ante* beliefs were wrong, *i.e.* did not correctly predict the historical risk-adjusted performance.

The methodology of this paper has the potential to discriminate between the two explanations because it allows market beliefs to be wrong (to a certain degree). The second explanation covers the case where the market has biased priors about the chances that options expire in the money but has otherwise correct beliefs. Modified returns or weighted modified

TABLE 2  
*Prices of CBOE S&P 500 index call options; risk adjustment based on Rubinstein's log-CAPM*

Selection bias	Return measure	<i>N</i>	Daily average (%)	Intercept	Slope
<i>All contracts</i>	Standard	1102	−0.925* (−1.767)	−0.031** (−5.526)	−0.954** (−3.455)
<i>Winning contracts</i>	Standard	760	1.102* (1.608)		
	Modified	760	−2.109** (−3.258)		
	Weighted standard	760	0.069** (3.506)		
	Weighted modified	760	−0.002 (−0.094)	−0.000 (−1.573)	−0.027** (−2.844)

*Remarks:* The table displays statistics computed from daily closing prices of SPX option contracts that were at the money 5 weeks before expiration. The sample covers the period 11 January 1991 until 7 December 1995. At all times, the nearest-maturity contract with between 5 and 1 week until expiration is followed. All prices are in real terms, converted from nominal prices using the monthly CPI numbers provided by the Federal Reserve Bank of Saint Louis (daily inflation rates are obtained by linear interpolation of monthly rates). Rubinstein's log-CAPM model is used to adjust for risk and time value of money (deflation). Modified rates of return are computed based on the *end-of-period* price; weighted rates of return are returns multiplied by the (deflated, real) payoff on the security at maturity. *N* is the number of observations. Intercept and Slope are for OLS projections of log returns onto moneyness (log ratio of strike price over index level). *t*-Statistics in parentheses (*t*-statistics of averages are adjusted for skewness as explained in the text). \*Significant at the 5% level; \*\*significant at the 1% level (both are one-sided significance levels).

returns should then be as predicted in Section 5.1. If the data confirm the predictions, the smile effect in performance is to be attributed to mistaken beliefs, and not to mistakes in empirical modelling of risk premia. That is, the first explanation is to be eliminated in favour of the second one.

Table 2 illustrates how the tests work for the S&P 500 options. Let us first look at the average daily rate of return computed in the usual way (daily deflated price change divided by previous day's deflated price).<sup>11</sup> We observe overpricing: Table 2 documents that the average risk-adjusted return is about −1% per day, and marginally significant.<sup>12</sup> This amounts to a weak rejection of EMH.

With the statistics from Section 5.1, we can investigate whether the (slight) mispricing of the S&P 500 options is attributable to biases in the market's prior about the chances that these options would expire in the money. There are two possible tests, both based on the modified rate of return of winning options (options that did expire in the money). First, there is the restriction in (16) that the modified rate of return should be negative on average. Second, there is the restriction in (17) that the mean weighted modified rate of return should be zero. The evidence in Table 2 confirms these predictions. The modified rate of return is significantly negative and the weighted modified rate of return is insignificantly different from zero.<sup>13</sup> To avoid problems of heteroscedasticity caused mainly by the increase in the S&P 500 index level over 1991–1995,

11. The time series of daily returns was pooled into one large sample. In Table 1, the data were manipulated differently: the time series mean was computed for each history, and cross-sectional averages were taken. The two procedures are equivalent.

12. Significance levels in Table 2 correct for skewness in the distribution of returns, as follows. Gross returns (one plus the rate of return) are assumed to be lognormally distributed. The logarithmic transformation is applied, and the hypothesis that returns equal one on average (*i.e.* rates of returns equal zero on average) is translated into the hypothesis that the logarithm of the returns equal minus one-half the variance in expectation. All *t*-tests are based on the latter hypothesis.

13. All *t*-statistics adjust for pronounced skewness in returns (modified, weighted modified). See previous footnote.

the weighted modified return is first divided by the strike price before computing the average. The insignificance of the weighted modified return is not to be attributed to the noise caused by weighting the modified rate of return with the deflated final payoff, because Table 2 shows that the weighted rate of return is significantly positive when the return is computed in the standard way. That is, if we use a closely related but inappropriate performance measure, we reject the null of zero.<sup>14</sup>

When we project the return on the options onto their moneyness, we recover a smile effect: overpricing (which translates into underperformance) becomes worse as the strike price increases relative to the index level. See the results in the first row of Table 2, in the columns “Intercept” and “Slope”. Shown are estimates and corresponding  $t$ -statistics for OLS projections of the logarithm of the return onto the moneyness, defined as the log ratio of strike price over index level.<sup>15</sup> Both the intercept and slope coefficient are significantly different from zero; the slope is negative.

This smile effect in performance cannot be explained in terms of Rubinstein’s log-CAPM and the belief model of the previous section. Table 2 documents that the weighted average modified rate of return for winners decreases as the strike price increases relative to the index level. To avoid problems of heteroscedasticity caused mainly by the increase in the S&P 500 index level over 1991–1995, the weighted modified return is first divided by the strike price before applying the OLS projection. The slope coefficient is still significantly negative.<sup>16</sup>

Evidently, we cannot discriminate between the two potential explanations for the smile effect, namely, mis-specified asset pricing model or mistaken beliefs. It turns out that the first explanation applies: our choice of Rubinstein’s log-CAPM is inappropriate. When we take Rubinstein’s more general model of power utility, with a coefficient of relative risk aversion equal to 5, modified returns and weighted modified returns exhibit the very patterns predicted in Section 5.1. Before this evidence is shown, let us first elaborate on how to incorporate preference parameters in the analysis.

## 6. INCLUDING PREFERENCE PARAMETERS

In the previous section, we took a simple but non-trivial asset pricing model to deflate prices and payoffs, namely, Rubinstein’s log-CAPM. The stochastic discount factor for this model, and hence, the price deflator, is parameter-free. In general, the deflator involves preference parameters (*e.g.* risk aversion coefficient). We first illustrate how to incorporate preference parameters in the analysis and then elaborate on how to estimate them along with a test of our theory.

Assume we choose Rubinstein’s general power utility asset pricing model to adjust for risk (his log-CAPM obtains for a special case, namely, logarithmic utility). In this model, the stochastic discount factor equals

$$A_t = \left( \frac{1}{\tilde{R}_t^M} \right)^\gamma, \quad (18)$$

all  $t$ .  $\gamma$  denotes the coefficient of risk aversion. This discount factor is applied to deflate prices and payoffs as in (5). Standard and modified (weighted) returns are then computed and the predictions in Section 5.1 are verified.

14. The standard and weighted standard rate of return is significantly positive, which is plainly the result of our selection bias: the statistics are computed only from prices of options that expired in the money.

15. The logarithmic transformation is applied to capture skewness in the return distributions.

16. There is a marked difference in magnitude between the standard return and the weighted modified return (divided by the strike price) which explains the differences in scale between the OLS projection results with the standard return as regressand (first row of Table 2) and those with the weighted modified return as regressand (last row of Table 2).  $t$ -statistics, however, are scale independent.

For instance, the prediction in Theorem 5 is

$$E \left[ \frac{W}{P_t} - \frac{W}{P_{t-1}} \middle| I_{t-1} \right] = 0.$$

Rewrite this in terms of stochastic discount factors and raw prices:

$$E \left[ \frac{A_{T^*} A_T A_{T-1} \dots A_{t+1} \tilde{W}}{\tilde{P}_t} - \frac{A_{T^*} A_T A_{T-1} \dots A_t \tilde{W}}{\tilde{P}_{t-1}} \middle| I_{t-1} \right] = 0.$$

Here,  $\tilde{P}_t$  denotes the raw (non-deflated) time- $t$  price, and  $\tilde{W}$  denotes the raw time- $T^*$  payoff. Substituting (18) for the stochastic discount factor. One obtains

$$E \left[ \left( \frac{1}{\tilde{R}_{t,T^*}^M} \right)^\gamma \frac{\tilde{W}}{\tilde{P}_t} - \left( \frac{1}{\tilde{R}_{t-1,T^*}^M} \right)^\gamma \frac{\tilde{W}}{\tilde{P}_{t-1}} \middle| I_{t-1} \right] = 0, \quad (19)$$

where  $R_{\tau,T^*}^M$  denotes the return on the market portfolio over the period  $(\tau, T^*)$  ( $\tau = t-1, t$ ).

As an illustration, consider again the S&P 500 option price histories discussed in the previous section. Re-evaluate the evidence using Rubinstein's general model instead of his log-CAPM to deflate prices and payoffs, setting  $\gamma = 5$ . Table 3 displays the results. It is a replication of Table 2 for Rubinstein's general model with  $\gamma = 5$ . The evidence against the EMH is now more pronounced because the significance level of the (negative) average standard return increases. The smile effect is still present: the slope coefficient in projections of the log standard return on the log ratio of strike price over index level is significantly negative. But the evidence against the belief model of Section 5.1 disappears: the slope coefficient in projections of the log weighted modified return of winners onto the log ratio of strike price over index level is now insignificant.<sup>17</sup> We can conclude that EMH is rejected not because we applied the wrong asset pricing model to adjust for risk, but because it makes assumptions about market beliefs that are too strong. Instead, the belief structure introduced in Section 5.1 matches the historical record more closely.

Table 3 displayed results for a particular preference parameter value only ( $\gamma = 5$ ). In general, one would like to find the preference parameter value that best fits the data and adjust significance levels for the corresponding search. Because the restrictions on weighted modified returns are cast in terms of moment conditions—see (19)—this is done most simply with the generalized method of moments (GMM). In particular, equation (19) can be used to generate a set of moment conditions. The distance of the sample version of the moments is minimized in order to obtain parameter estimates. The distance is measured as the weighted sum of squares. If an estimate of the asymptotic covariance matrix of the moments is used as weighting matrix, the minimum distance is asymptotically  $\chi^2$ -distributed. See Hansen and Singleton (1982).

## 7. RESTRICTION ON LOSERS

Until now, implications of our belief structure have been shown only for prices of “winning” securities. One obviously wonders whether restrictions obtain for prices of “losing” securities

17. To avoid problems of heteroscedasticity, the weighted modified return is first divided by the strike price before applying the OLS projection. The intercept is small but significantly negative, which is consistent with the theory, provided returns are lognormally distributed. From Theorem 5, the weighted modified rate of return should be zero on average even after division with the strike price. Adding one and taking logarithms then produces a variable with an expected return equal to minus one-half its variance, a small, negative quantity (when divided by the strike price, it is approximately  $-10^{-5}$ ). Under the hypothesis that the slope coefficient in the projection onto moneyness is zero, the intercept should therefore be significantly negative.

TABLE 3

*Prices of CBOE S&P 500 index call options; risk adjustment based on Rubinstein's general power utility model with relative risk aversion equal to 5*

Selection bias	Return measure	<i>N</i>	Daily average (%)	Intercept	Slope
<i>All contracts</i>	Standard	1102	-1.407** (-2.897)	-0.033** (-6.439)	-0.954** (-3.830)
<i>Winning contracts</i>	Standard	760	0.380 (0.586)		
	Modified	760	-2.168** (-3.759)		
	Weighted standard	760	0.028** (2.135)		
	Weighted modified	760	-0.019 (-1.545)	-0.000** (-2.162)	-0.010 (-1.561)

*Remarks:* The table displays statistics computed from daily closing prices of SPX option contracts that were at the money 5 weeks before expiration. The sample covers the period 11 January 1991 until 7 December 1995. At all times, the nearest-maturity contract with between 5 and 1 week until expiration is followed. All prices are in real terms, converted from nominal prices using the monthly CPI numbers provided by the Federal Reserve Bank of Saint Louis (daily inflation rates are obtained by linear interpolation of monthly rates). Rubinstein's general power utility model is used to adjust for risk and time value of money (deflation), with coefficient of relative risk aversion set equal to 5. Modified rates of return are computed based on the *end-of-period* price; weighted rates of return are returns multiplied by the (deflated, real) payoff on the security at maturity. *N* is the number of observations. Intercept and Slope are for OLS projections of log returns onto moneyness (log ratio of strike price over index level). *t*-Statistics in parentheses (*t*-statistics of averages are adjusted for skewness as explained in the text).

\*Significant at the 5% level; \*\*significant at the 1% level (both are one-sided significance levels).

(*i.e.* those that end in the default state). Subject to conditions on the nature of the prior of the market, the answer is affirmative. However, restrictions obtain only when investigating the return series in reverse time. They again follow from a general result about the evolution of Bayesian posteriors. Now, however, the prior has to be random. More specifically, the distribution of priors has to be non-informative.

The analysis is complex, not only because time is to be reversed, but also because the notion of a non-informative prior is subtle. Therefore, the analysis is not included here. The interested reader may consult Bossaerts (1996).

## 8. CONCLUSION

This article essentially showed that the key assumption behind much of empirical finance, namely, the EMH, is unnecessarily strong. It derived a novel property for Bayesian posteriors, and from it, a number of statistics that can be used to test asset pricing theory on limited-liability securities prices when the market is not assumed to know the correct probability of the default state. Still, the market is required to update its beliefs on the basis of the correct likelihood function (of signals given the future payout).

This article's empirical illustrations, as well as the empirical tests in Bondarenko (1998), Bondarenko and Bossaerts (2000) and Bossaerts and Hillion (2001) demonstrate that the new belief model is capable of capturing some salient features of historical returns when EMH cannot. In the process, only simple asset pricing models appear to be required in order to account for risk premia, such as Rubinstein's log-CAPM or his more general power utility model. This counters the recent tendency to appeal to increasingly complicated asset pricing models (*i.e.* stochastic discount factors) to explain the historical record.

It deserves emphasis that the ability of our belief model to better explain historical data cannot be attributed to our allowing for more degrees of freedom. Inspection of the statistics

reveals that the number of parameters does not increase. When no parameters appear in tests of an asset pricing model under EMH (*e.g.* risk neutrality—used in Section 4.2; Rubinstein’s log-CAPM—used in Section 5.2), then none are present under our belief model. When parameters appear under EMH (*e.g.* Rubinstein’s general power utility model with power utility—used in Section 6), then the same number are present under our belief model. Hence, our statistics do not relate to standard EMH-based statistics through the addition of parameters. Instead, they obtain through deliberate selection bias, and through manipulation of price data into non-standard return measures. Admittedly, because of the deliberate selection biases, data will have to be thrown out, potentially reducing the power of tests of our belief model compared with tests of EMH. But the selection biases often occur naturally, such as when the empiricist only observes the performance of winners, in which case tests of EMH are impossible anyway unless one is willing to formulate hypotheses about the features of the non-observed sample.

The success of the belief model indicates that it may be more fruitful to attempt to explain why market priors often appear to be biased. This may not just reflect historical accident, but perhaps psychological biasing in the face of ambiguity aversion, as in the multiple-prior context of Gilboa and Schmeidler (1989).<sup>18</sup> If one interprets the parameter chosen by the maximin agent of Cagetti *et al.* (2000) and Maenhout (2001) as a prior, biases emerge for robustness reasons.

Still, the term “belief bias” has to be handled with care. If one discovers that market priors were biased, it is not clear whether this means that the market “gets it wrong”. It only implies that there is a discrepancy between the market’s prior belief about security payoffs and the historically recorded frequency distribution. The latter may itself be biased if, for instance, the empiricist collected a biased sample. Or it may not have any implication for future payoffs if the payoff generating process is non-stationary.

Several topics for future research emerge. First, if one rejects an asset pricing model with the standard EMH belief structure, but does not with our model of beliefs, the evidence points to biases in market priors. An example is in Section 6, where S&P 500 option prices in the first half of the 1990s were shown not to conform to EMH, but instead to our belief model, along with Rubinstein’s general power utility model with coefficient of relative risk aversion equal to 5. One wonders whether a trading rule could be derived that exploits such a finding. The answer is related to an earlier comment. Our statistics are not based on estimation of a number of parameters that would not be required in standard EMH-based methodology. In particular, one does not estimate the market’s initial beliefs at the beginning of each history. Any trading rule, however, would have to be based on information on how the market’s beliefs are biased. That is, it would require one to estimate the very parameters that the tests in this article avoid. And they avoid it for a good reason: if a theory can only be tested by adding parameters, it may fit the data too easily.

Second, imagine that the payoff generating process is stationary, and that the empiricist studies the market’s pricing over a period  $t = 1, 2, \dots, T$ . In particular, the empiricist attempts to identify any predictability in (properly deflated) returns. Once this is done, the empiricist uses the model so identified to predict (properly deflated) returns out of sample, *i.e.* over a period  $t = T + 1, T + 2, \dots, T^*$ . Will the empiricist be successful? In particular, if the empiricist projects out-of-sample (deflated) returns onto his/her forecasts, will the slope coefficient turn out to be significant a higher proportion than expected if it were impossible to predict (deflated) returns? What about the intercept? The answers are not obvious. On the one hand, the market is implicitly assumed to be Bayesian, and a Bayesian may happen to have started with the “right”

18. The arguments that Daniel and Titman (1999) use to explain overconfidence in stock markets also involve ambiguity. Instead of the ambiguity aversion of the formal literature—*e.g.* Gilboa and Schmeidler (1989)—Daniel and Titman seem to invoke ambiguity seeking.

prior, in which case it would be hard to predict future mistakes. Or the market may have started with the “wrong”, even obstinate priors, in which case it may be easy to predict mistakes. Some answers can be found in a working paper, Bossaerts (1997).

Third, the cross-sectional implications of the approach of this paper need to be studied. Under the standard EMH-based methodology, the realized average return only reflects compensation for risk (modulo sampling error). In cross-section, this means that securities with a higher average return must somehow have been more risky (asset pricing theory is given the task to identify the risk). In the model of this article, a security’s *ex post* average return also generally reflects news on the basis of which the market updated a biased *ex ante* belief. That is, *ex post* average returns in part reflect belief biases. For example, MS’s average stock return over the last 15 years has been high not necessarily because it was a high risk company, but probably mostly because the market initially thought it to be an average company, and subsequently had to revise its belief as evidence mounted of exceptional profit growth. It is clear that such a view of the world may shed new light on extant cross-sectional puzzles, and hence, it deserves further research.<sup>19</sup>

*Acknowledgements.* The reader who is interested in a verbose version of this paper, with extensive numerical examples, is referred to Bossaerts (1998). The first part of this paper came out of a sketchy set of notes that the author presented at the Université des Sciences Sociales, Toulouse, in April 1995, while he was at CentER, Tilburg University. He is grateful for the many comments from the seminar participants. In addition, the paper benefited through discussions at seminars at Carnegie Mellon University, U.C. Berkeley, U.C.L.A., U.C. Riverside, University of Minnesota, Washington University at Saint Louis, the 1996 Asset Pricing summer meeting at NBER, the 1996 European Meetings of the Econometric Society, the 1998 CEPR Meetings in Gerzensee, and through comments from and discussions with Oleg Bondarenko, James Dow, Darrell Duffie, David Easley, Ravi Jagannathan, Pierre Hillion, Alan Kraus, Guy Laroque, P.C.B. Phillips, Richard Roll, Philippe Weil and three anonymous referees. Oleg Bondarenko provided able research assistance. The usual disclaimer applies.

#### REFERENCES

- BONDARENKO, O. (1998), “Testing Rationality of Financial Markets: An Application to S&P 500 Index Options” (Ph.D. Dissertation, Caltech, Division of Humanities and Social Sciences).
- BONDARENKO, O. (2001), “Why are Put Options so Expensive?” (UIC Finance Department Working Paper).
- BONDARENKO, O. and BOSSAERTS, P. (2000), “Expectations and Learning in Iowa”, *Journal of Banking and Finance*, **24**, 1535–1555.
- BOSSAERTS, P. (1996), “Martingale Restrictions on Securities Prices Under Rational Expectations and Consistent Beliefs” (Caltech Working Paper).
- BOSSAERTS, P. (1997), “A Theorem on (Certain Kinds of) Out-of-sample Prediction Tests in Finance” (Caltech Working Paper).
- BOSSAERTS, P. (1998), “The Dynamics of Equity Prices in Fallible Markets” (Caltech Working Paper).
- BOSSAERTS, P. (1999), “Learning-Induced Securities Price Volatility” (Caltech Working Paper).
- BOSSAERTS, P. and HILLION, P. (2001), “IPO Post-Issue Markets: Questionable Predilections but Diligent Learners?”, *Review of Economics and Statistics*, **83**, 333–347.
- CAGETTI, M., HANSEN, L. P., SARGENT, T. and WILLIAMS, N. (2000), “Robustness and Pricing with Uncertain Growth” (Working Paper, University of Chicago, Department of Economics).
- DANIEL, K. and TITMAN, S. (1999), “Market Efficiency in an Irrational World”, *Financial Analysts Journal*, **55**, 28–40.
- DE LONG, B., SHLEIFER, A., SUMMERS, L. and WALDMAN, R. J. (1990), “Positive Feedback Investment Strategies and Destabilizing Rational Speculation”, *Journal of Finance*, **45**, 375–395.
- DIACONIS, P. and FREEDMAN, D. (1986), “On the Inconsistency of Bayes Estimates”, *Annals of Statistics*, **14**, 1–67.
- DOOB, J. L. (1948), “Application of the Theory of Martingales”, *Colloques Internationaux du CNRS*.
- FAMA, E. (1970), “Efficient Capital Markets: A Review of Theory and Empirical Work”, *Journal of Finance*, **25**, 383–417.
- FAMA, E. F. (1991), “Efficient Capital Markets: II”, *Journal of Finance*, **46**, 1575–1618.
- GILBOA, I. and SCHMEIDLER, D. (1989), “Maxmin Expected Utility with Non-unique Priors”, *Journal of Mathematical Economics*, **59**, 33–49.
- HANSEN, L. P. and SINGLETON, K. J. (1982), “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models”, *Econometrica*, **50**, 1269–1286.

19. Daniel and Titman (1999) explore this path, by arguing that the ambiguity in a firm’s environment may cause biases in *ex ante* beliefs which translate into differences in *ex post* performance relative to other firms.

- LEWELLEN, J. and SHANKEN, J. (2000), "Estimation Risk, Market Efficiency, and the Predictability of Returns" (NBER Working Paper).
- LIU, J. and LONGSTAFF, F. (2000), "Losing Money on Arbitrages: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities" (UCLA Anderson Graduate School of Management Working Paper).
- LUCAS, R. (1978), "Asset Prices in an Exchange Economy", *Econometrica*, **46**, 1429–1445.
- MAENHOUT, P. (2001), "Robust Portfolio Rules, Hedging and Asset Pricing" (INSEAD Working Paper).
- MEHRA and PRESCOTTE. (1985), "The Equity Premium: A Puzzle", *Journal of Monetary Economics*, **15**, 145–161.
- RUBINSTEIN, M. (1976), "The Valuation of Uncertain Income Streams and the Pricing of Options", *The Bell Journal of Economics*, **7**, 407–425.
- SMITH, L. (1996), "Private Information and Trade Timing" (MIT Working Paper).